8/PRTS
09/701572
528 Rec'd PCT/PTO 08 DEC 2000
WO 99/64451
PCT/FR99/01342

PLANT PROTEIN WITH REPEATED WD40 MOTIFS, NUCLEIC ACID CODING FOR SAID PROTEIN, AND USES THEREOF

The invention relates to the cloning of genes involved in regulating cell division in plants, and their uses.

Most plant organs develop after germination, through differentiation from the meristems. Prior to differentiation, the cell division cycle slows down and then stops in the meristems. Simultaneously, an increase in the size of the cells, and replication of the genome not accompanied by mytosis, called "endoreplication", are frequently observed. Endoreplication is a well known phenomenon during the development of storage tissue; KOWLES [Genome, 35, pp. 68-77, (1992)] thus mention a ploidy of 6C to 384C during the development of the endosperm in maize.

The phenomena involved in the stoppage of cell division preceding differentiation play an essential role in plant development and ontogeny. The mechanisms involved in these phenomena are still poorly known; it appears that the inhibition of the factor for promoting the M phase, and the induction of the protein kinases of the S phase (GRAFI, Science, 269, pp. 1262-1264, (1995)] could be involved. However, no factors directly involved in this mechanism have so far been identified in plants.

The inventors undertook the study of this mechanism with the aim of discovering the means of controlling and of acting thereby on plant development and ontogenesis.

They chose, as a study model, the *Rhizobium*/leguminous plant symbiotic system. In this system, the Nod factors, which are lipooligosaccharide in nature and which are produced by *Rhizobium*, constitute mitogenic signals which locally induce the formation of a new meristem, from which the cells forming the root nodules become differentiated [TRUCHET, Nature, 351, pp. 670-673, (1991); YANG, Plant Cell, 6, pp. 1415-1426, (1994); SAVOURE, EMBO.J., 13,

09701572.120801

pp. 1093-1102, (1994)]. The nodules comprise 3 main regions: an apical region, consisting of meristematic cells; an intermediate region for invasion or for differentiation (region II), where the infection of the cells by bacteria, as well as the stoppage of cell division, accompanied by endoreplication and an increase in the size of the cells, followed by their differentiation, take place; and a region for fixation (region III), consisting of differentiated cells infected by bacteria, and where the fixation of nitrogen takes place.

During this study, the inventors isolated, from lucerne *(Medicago sativa)* nodules, a gene, called hereinafter *ccs52*, which plays an essential role in the stoppage of the cell cycle and the induction of endoreplication. Using a cDNA probe of the *Medicago sativa ccs52* gene, they also isolated a homologous gene in *Medicago truncatula*.

The *ccs52* genes of *Medicago sativa (ccs52Ms)*, and of *Medicago truncatula (ccs52Mt)* encode a polypeptide of 475 amino acids having a theoretical molecular mass of 52 kDa. These polypeptides are called hereinafter CCS52Ms and CCS52Mt, respectively; the sequences of CCS52Ms and CCS52Mt differ by only 2 residues at positions 16 (R/G) and 141 (V/I).

These 2 proteins comprise repeated WD motifs, and may thus be attached to the superfamily of proteins with repeated WD motifs.

The repeated WD motifs comprise about 40 amino acids containing a number of conserved amino acids including the WD motif (Trp-Asp) which is frequently situated at one end of the repeated motif [NEER et al., Nature, 371, pp. 297-300, (1994)]. The members of this family regulate various functions, such as signal transduction, transcription, pre-mRNA splicing, organization of the cytoskeleton, vesicular fusion or the cell cycle. Although the general structure is overall similar in all the proteins, the wide functional variety of repeated WD motifs suggests that

these motifs have become differentiated and have become functionally specialized. A functional homology is reflected in the number of repeated WD motifs, by a strong homology of the repeated WD motifs with equivalent positions in various proteins, compared with other repeated motifs in the same proteins, and by a significant similarity of the C- and N- terminal ends.

Comparison of the sequence of CCS52Ms with the sequences of known proteins, using the GENETICS COMPUTER GROUP GAP programme [parameters: gap weight: 1000; length weight: 0.100; average match: 0.540; average mismatch: 0.396] reveals a high homology with the proteins containing repeated WD40 motifs which are involved in the regulation of the cell cycle, and more specifically, with the Drosophila FZR proteins (57% identity), *Saccharomyces cerevisiae* HCT1 (46% identity), and *Schizosaccharomyces pombe* SRW1 (52% identity), which belong to the "fizzy-related" (FZR) family. Research carried out on databases of sequences using the BLAST programme [ALTSCHUL et al. Nucleic Acids Res. 25:3389-3402, (1997)] have also shown a strong homology of CCS52Ms with the Drosophila FZR proteins (56% identity; 70% similarity), and the *Schizosaccharomyces pombe* SRW1 proteins (51% identity; 67% similarity) mentioned above, as well as with the product of the *X. laevis fzr* gene (58% identity; 73% similarity).

The FZR proteins induce the degradation of the mitotic cyclins and are involved in the transition between cell proliferation and differentiation. It has thus been shown in Drosophila that the *fzr* gene is expressed at the end of cell proliferation during embryogenesis. The product of this gene causes a reduction in the mitotic cyclins, and is necessary for the stoppage of cell proliferation and the start of the endocycles [SIGRIST and LEHNER, Cell, 90, pp. 671-681, (1997)]. In *Saccharomyces cerevisiae*, HCT1 is necessary for the proteolysis of the mitotic cyclin, Clb2 [SCHWAB et al., Cell, 90, pp. 683-693, (1997)]. In

Schizosaccharomyces pombe, the product of the *swr1* gene controls the cell cycle and differentiation by negatively regulating the Cdc2/CDC13 (cyclin of the mitotic type) complexes [YAMAGUCHI et al., Mol. Biol. Cell., 8, 2475-2486, (1997)]. The FZR proteins therefore have a different role from that of the other proteins with repeated WD motifs, which are involved in cell proliferation.

In plants, no protein of the FZR family had been described prior to CCS52Ms.

The existence of a gene encoding a protein with repeated WD40 motifs and its isolation from carrot cDNA have recently been described [LUO et al., Plant Mol. Biol., 34, pp. 325-330, (1997)]. However, the product of this gene exhibits a weaker homology (44% identity and 63% similarity on the sequence comparison carried out with the BLAST programme) with the CCS52Ms protein than the FZR proteins of invertebrates and of yeast; this carrot protein is related to the cdc20, p55 and fizzy proteins, and therefore belongs to a subgroup of proteins with repeated WD40 motifs distinct from the FZR subgroup.

The search for homologues of CCS52Ms in a database of the *Arabidopsis thaliana* genome has revealed a peptide sequence deduced from a genomic clone (AB005230) and exhibiting 64% identity with CCS52Ms, which shows the existence of homologues of the *ccs52Ms* gene in other plants. Another peptide sequence also deduced from a genomic clone of *Arabidopsis thaliana* (AL031018, published on 17 September 1998) exhibits 80% identity with CCS52Ms (44% identity and 63% similarity based on the sequence comparison carried out with the BLAST programme).

Figure 1A represents a dendrogram of the family of proteins with repeated WD40 motifs, which shows that the CCS52 proteins form with the other FZR proteins a subfamily representing a branch which evolved separately from those respectively consisting of the CDC20, P55 and fizzy proteins.

Sub B1

Figures 1B and 1C represent the alignment, carried out using the "PRETTYBOX" software, of the *Medicago sativa* CCS52 (MsCCS52) sequence and the Drosophila FZY and FZR (DmFZY and DmFZR) sequences, of the *Saccharomyces cerevisiae* HCT1 (ScHCT1) sequence, the *Schizosaccharomyces pombe* SRW1 (SpSRW1) sequence, the *Arabidopsis thaliana* FZY (AtFZY) sequence and the 2 *Arabidopsis thaliana* polypeptides (AtCCS52A = peptide deduced from AL031018, and AtCCS52B = peptide deduced ~~from AB~~005230).

The CCS52Ms protein contains 7 domains with repeated WD40 motifs, situated in the central and C-terminal portions of the molecule (the location of these domains numbered from I to VII, is indicated in Figures 1B and 1C, above the alignment of the sequences). These domains exhibit only a slight homology with each other, hence it can be concluded that they represent sites for interaction with different proteins. The latter domain (VII) comprises a potential binding site for the cyclins.

In the N-terminal portion of the CCS52Ms protein are localized a peptide sequence (DRFIPSR) which corresponds to a motif present in the FZR proteins as well as in other proteins with repeated WD40 motifs such as cdc20, p55 and fizzy, as well as a peptide sequence (AYTTLLRTALFG) which corresponds to a motif specific to the FZR family, absent from the other proteins with repeated WD40 motifs (the location of these motifs, called I and II respectively, is indicated in Figure 1B above the alignment of the sequences).

Potential sites for phosphorylation with CDKs (cyclin-dependent kinases) are located in the N-terminal portion, at positions 43 (SPSR), 99 (TPEK), 144 (SPVK), 154 (RSP) and 155 (SPYK), as well as in the C-terminal portion at position 454 (SPK), of CCS52Ms. The sites situated at positions 43 and 144 are also present in other FZR proteins, whereas the sites situated at positions 99, 154 and 155 appear more

specific to the CCS52 proteins of plants; the C-terminal site at position 454 also appears to be specific to the CCS52 proteins of plants.

A sequence of 15 amino acids RDNSPPPEPSPESLR starting at residue 16, and corresponding to a protein degradation motif PEST is also present in the N-terminal portion of CCS52Ms. This motif probably makes it possible, through the degradation of CCS52, to regulate its interactions with other proteins.

The structure of the CCS52Ms protein is schematically represented in Figure 2, in which the position of the WD40 motifs, of the phosphorylation sites (P), of the PEST motif, and of the I and II motifs, are indicated.

The sequence of the *Medicago sativa* cDNA cloned by the inventors is represented in the sequence listing in the annexe under the number SEQ ID NO:1; the sequence of the corresponding CCS52Ms protein is represented under the number SEQ ID NO:2.

The untranslated 3′ region of the transcript of this DNA comprises 2 AUUUA sequences, which correspond to sequences for instability of the mRNA, and may therefore play a role in regulating the quantity of transcripts of *ccs52*.

The inventors searched for the presence of homologues of *ccs52Ms* by Southern transfer, in diploid and tetraploid species of *Medicago*, as well as in other plants, in particular tobacco, tomato, potato, soya, wheat and rice: in all cases, several bands were detected, which indicates that *ccs52* indeed represents a family of plant genes which is related to the *fzr* family.

The inventors studied *in vivo* the activity of the CCS52Ms protein and showed that it was involved in regulating cell differentiation, and in promoting endoreplication. In particular, the expression of the CCS52Ms protein in transgenic plants induces therein an increase in endoreplication and in the level of ploidy of the cells of plants. This effect could be the

consequence of a blocking of mitosis by the activation of the degradation of the mitotic cyclins, which would bring about conversion of the mitotic cycles to endocycles consisting of the G1-S-G2 phases. The result of the repetition of the endocycles is the amplification of the genome and the increase in ploidy, correlated with an increase in cell volume.

The subject of the present invention is a plant protein with repeated WD40 motifs, called CCS52, characterized in that it belongs to the FZR subfamily.

According to a preferred embodiment of the present invention, the said plant protein exhibits at least 45%, and preferably at least 55% identity with the polypeptide having the sequence SEQ ID NO:2 or at least 60% and preferably at least 70% similarity with the polypeptide having the sequence SEQ ID NO:2.

The present invention includes in particular the CCS52Ms protein, its isoforms, as well as the autologous proteins of *Medicago* and the orthologous proteins of other plants, which may be attached to the family of FZR proteins.

The invention also includes proteins derived from the CCS52 proteins by addition, deletion or substitution of one or more amino acids or of one or more amino acid sequences; this may include for example proteins in which modifications have been made outside the functional regions, or alternatively proteins in which modifications have been made in order to modify their activity, for example proteins stabilized by deletion of the PEST motif.

The subject of the present invention is also a purified nucleic acid fragment, characterized in that it comprises all or part of a sequence encoding a CCS52 protein, as defined above, or its complementary sequence. In this context, the present invention includes in particular the cDNAs and the genomic DNAs of the CCS52 proteins.

Nucleic acid fragments in accordance with the present invention can be easily identified and cloned

by screening plant cDNA or genomic DNA libraries with the aid of oligonucleotides derived from the *ccs52Ms* sequence, and in particular oligonucleotides derived from the regions of this sequence which are specific to the FZR proteins, and in particular the CCS52 proteins.

The CCS52 proteins may be produced, in particular, by expressing these nucleic acid sequences in host cells.

The subject of the present invention is also the use of a CCS52 protein, as defined above, or of a nucleic acid sequence encoding all or part of the said protein, or of its complementary sequence, for regulating the differentiation and the proliferation of plant cells.

The subject of the present invention is also the use of a protein of the FZR subfamily or of a nucleic acid sequence encoding all or part of the said protein, or of its complementary sequence, for regulating the differentiation and the proliferation of plant cells.

There may be mentioned, among such proteins, the drosophila FZR protein or the yeast FZR protein.

The modification of the expression and/or of the activity of CCS52 proteins in plant cells makes it possible to modify the cell cycle, by promoting either proliferation or differentiation, and to thus control the development process, in order to obtain, for example, stimulation of somatic embryogenesis, to increase *in vitro* regeneration of plants from calli, by increasing conversion to embryos, or to promote the development of certain organs, for example to increase the productivity of storage tissues by increasing their endoploidy.

It is possible in particular to use the cDNA sequences of CCS52 proteins or of portions of these cDNA sequences, or of their sense or antisense transcripts; this may be for example the entire sequence encoding a CCS52Ms protein, or a portion of this coding sequence, and/or all or part of the

untranslated 5′ and 3′ regions. These sequences may be used in the sense orientation, or if it is desired to inhibit the expression of the CCS52Ms protein in a plant or in a tissue or organ thereof, in antisense orientation.

The present invention also includes recombinant DNA constructs containing at least one nucleic acid sequence in accordance with the invention.

Generally, the said nucleic acid sequence will be placed under transcriptional control of an appropriate promoter.

Advantageously, it will thus be possible to use a strong promoter in order to increase, in the host cells, the levels of expression of the CCS52 protein; this may include an inducible promoter or a constitutive promoter, a ubiquitous promoter, or a tissue-specific promoter.

The use of inducible promoters makes it possible to obtain blocking of mitosis, and the induction of endoreplication at the desired moment. The use of tissue-specific promoters makes it possible to target the action of the CCS52 protein at certain tissues and organs (for example storage tissues).

By way of examples of strong promoters which can be used in the context of the present invention, there may be mentioned: the CaMV35S [BENFLY et al., Science, 250, pp. 959-966, (1990)], the 35S promoter; the *Agrobacterium tumefaciens* T-DNA promoters: nopaline synthase, octopine synthase, mannopine synthase, 1′, 2′ [SANDERS et al., Nucleic Acid Res., 15, pp. 1543-1558, (1987); HOOYKAAS and SCHILPEROORT, Plant. Mol. Biol., 19, pp. 15-38, (1992)].

By way of examples of inducible promoters which can be used in the context of the present invention, there may be mentioned: the promoter inducible by tetracycline [WEINMANN et al., Plant J., 5, pp. 559-569, (1994)]; the promoter inducible by copper [METT et al., Transgenic Res., 5, pp. 105-113, (1996)]; the

promoter inducible by glucocorticoids [AOYAMA and CHUA, Plant. J., 11, pp. 605-612, (1997)].

By way of examples of tissue-specific promoters which can be used in the context of the present invention, there may be mentioned: the endosperm-specific promoter [OPSAHL-FERSTAD et al., Plant J., 12, pp. 235-246, (1997); DOAN et al., Plant Mol. Biol., 31, pp. 877-886, (1996); the nodule-specific promoters (*enod12A/B* or leghaemoglobin) [TRINH et al., Plant Cell Reports, (17, pp. 345-355, (1998); VIJN et al., Plant Mol. Biol., 28, pp. 1103-1110, (1995)] or early promoters inducible by the Nod factor and late promoters (promoter of cyclin D or of late nodulins (leghaemoglobin type) and promoters regulated by hormones, such as *parA/B* [TAKAHASHI et al., Proc. Natl. Acad. Sci, USA, 87, pp. 8013-8016, (1990)], *GH3* [LIU et al., Plant Cell, 6, pp. 645-657, (1994)].

The invention includes in particular recombinant vectors carrying at least one insert containing a DNA fragment in accordance with the invention. These vectors can be used for transforming host cells.

The subject of the invention is also cells and pluricellular organisms transformed with at least one DNA sequence in accordance with the invention; this includes in particular plant cells or plants.

The present invention will be understood more clearly with the aid of the additional description which follows, and which refers to nonlimiting examples illustrating the identification, cloning and expression of the CCS52Ms gene.

EXAMPLE 1: CLONING AND SEQUENCING OF A CCS52MS cDNA

A cDNA clone of CCS52Ms was obtained by differential screening from a cDNA library of *Medicago sativa* nodules, highly stimulated during nodular organogenesis.

The following protocol was used:

The cDNA of *M. sativa ccs52Ms* was isolated by the DD-RT-PCR (Differential Display RT-PCR) technique

[LIANG and PARDEE, Science, 257, pp. 967-971, (1992)], using the RNAimage® kits (GENHUNTER CORPORATION). The RNA samples are isolated from the root region sensitive to the Nod factor of young *M. sativa* plants (growth in a nitrate-limited medium), in the absence of bacteria or inoculated with Nod$^+$ (EK1433) or Nod$^-$ (EK133) strains of *R. meliloti* for 4 days. The DD-RT-PCR *ccs52Ms* fragment, exhibiting an increase in the expression of the nodules, is cloned into the cloning vector pCT-TRAP (GENHUNTER CORPORATION) and used as a probe for the isolation of complete clones from a cDNA library of nodules of *M. sativa sp. varia* A2, constructed in λ-ZAP (STRATAGENE) (CRESPI et al., EMBO J., 1994, 13, 5099-5112).

Seven cDNA clones, obtained from 2.10^5 phages, represent 2 types of cDNA differing from each other only in the 4 amino acids (16R-G, 17D-N, 33S-N, 52R-G) and the length of the 3'UTR fragment. A 99% identity for the clones, at the level of the amino acid sequence, suggests that they represent allels of the same gene in allogamous tetraploid *M. sativa*.

The sequencing of the *ccs52Ms* cDNA is carried out with the PERKIN-ELMER ABIprism system.

The genomic clones *ccs52Ms* and *ccs52Mt* are isolated from genomic libraries of *M. sativa* cv. Nagyszénasi and *M. trucatula* ecotype GHOR, using the *ccs52Ms* cDNA as hybridization probe. These genomic libraries are constructed by partial digestion of the genomic DNA with the restriction enzyme MboI and the cloning of the DNA fragments having a size of between 15 and 20 Kb into the BamHI site of λ-EMBL4.

EXAMPLE 2: IDENTIFICATION OF THE FAMILY OF THE *CCS52MS* GENE IN *MEDICAGO* AND ITS EXPRESSION IN VARIOUS PLANT ORGANS

The existence of multiple copies of the *ccs52* gene is tested for by hybridization of the Southern type in tetraploid cultivars of *M. sativa* Nagyszénasi and Cardinal and in autogamous diploid *M. truncatula*, a model plant in research on vegetables.

The plant DNA is isolated from young leaves, using the NUCLEON PHYTOPURE DNA extraction kit (AMERSHAM).

The DNA samples are digested with EcoRI and transferred onto BIOTRANS nylon membrane (+) (ICN).

The Southern hybridization is carried out in accordance with conventional protocols [(SAMBROOK, Molecular Cloning: A Laboratory Manual 2nd edn., Cold Spring Harbor Laboratory Press, New York, (1989); AUSUBEL, Current Protocols in Molecular Biology, (1989)], under stringent conditions at 65°C (hybridization in CG buffer; washing: 2 x SSC, 0.1% SDS for twice 15 min, then 0.5 x SSC, 0.1% SDS for twice 30 min).

The expression of *ccs52Ms* is studied by Northern analysis.

Total RNA is isolated from various organs of *M. sativa* cultivar Sitel:

- from the roots, inoculated for 4 days with the *R. meliloti* Nod$^-$ mutant (EK133) and with the strain overproducing Nod factors (EK1433);

- from the nodules, 12, 19, 23 and 30 days after infection with *R. meliloti*, and

- from the stems, hypocotyls, leaves, buds, flowers, roots of plants which are 3 days old, 7 days old, roots deprived of nitrogen and which do not have root tips, roots which are 7 days old, without root tips, placed in culture in the presence of nitrate, spontaneous nodules developed in the absence of *R. meliloti*, and root tips or a culture of cells of *M. sativa sp. varia* A2.

100 mg of each of the organs tested, collected under liquid nitrogen, are used for the extraction of the RNA (RNEASY PLANT, QUIAGEN).

The RNA is loaded (10 µg per lane) onto a denaturing gel (formaldehyde) [SAMBROOK, Molecular Cloning: A Laboratory Manual 2nd edn., Cold Spring Harbor Laboratory Press, New York, (1989)].

The DNA is transferred into a 10 x SSC transfer solution [CHOMCZYNSKI et al., Analytical Biochemistry, 221, pp. 303-305, (1994)].

Both in the case of the Southern hybridization and in the case of the Northern hybridization, the *ccs52Ms* cDNA fragment is labelled with [α^{32}P]dCTP (kit MEGAPRIM, AMERSHAM). Hybridization with the *Msc27* probe serves as control for the loading of the RNA [SAVOURE et al., EMBO J., 13, pp. 1093-1102, (1994)].

The results of the Southern transfer show that the probe hybridizes with various EcoRI fragments of the genomic DNA of *M. sativa* or *M. truncatula*, which indicates that *ccs52Ms* represents, in *Medicago*, a multigene family.

The results of the Northern transfer obtained with the total RNA of roots inoculated with the Nod$^-$ EK133 mutant of *R. meliloti*, or with the EK1433 strain overproducing Nod factors and with the RNA extracted from the nodules, 12, 19, 23 and 30 days after infection with *R. meliloti* show that only a small quantity of transcripts is observed in the total RNA of the roots, which reflects the small proportion of cells involved in the organogenesis of the nodules compared with the total number of cells of the roots. By contrast, in the nodules of different ages, a high level of transcription is observed, which reflects the persistence of the apical meristems and of the regions for differentiation.

The results of the Northern transfer which are obtained with the total RNAs of: 1: culture of cells of *M. sativa* sp. *varia* A2, 2: stems, 3: hypocotyls, 4: leaves, 5: flower buds, 6: flowers, 7: roots of shoots which are 3 days old, 8: roots of shoots which are 7 days old, deprived of nitrogen, lacking ends, 9: root tips which are 7 days old, cultured in the presence of nitrates, lacking ends, 10: spontaneous nodules developed in the absence of *R. melioti*, 11: nitrogen-fixing nodules, 12: ends of root tips, show that the expression of *ccs52Ms* is not limited to the nodules,

although this organ is that which contains the highest level of transcripts.

These transcripts are indeed present in variable quantities practically in all the organs, which indicates that this protein is involved in the development of each of them. Apart from the nodules, the level of transcription is also high in young shoots, and, in cell cultures, where a smaller sized mRNA is in addition detected which may correspond either to a different polyadenylation, or to the expression of a homologous copy of the gene.

Analyses were also carried out by *in situ* hybridization, and show that the mRNA of *ccs52Ms* is located mainly in the region for differentiation, and in particular at the interface between regions II and III of the nodule, which are regions where differentiation is the most active.

In parallel, expression of the G1 and mitotic type cyclins as well as of the H3 histone specific to the S phase is observed in the same regions.

This indicates that CCS52Ms is involved in the regulation of the cell cycle, probably in a manner similar to its yeast and drosophila homologues, that is to say by means of the proteolysis of mitotic cyclins, which inhibits mitosis and induces endoreplication cycles.

EXAMPLE 3: EXPRESSION OF CCS52MS IN SCHIZOSACCHAROMYCES POMBE

The expression of CCS52Ms was studied in *S. pombe* in which a functional homologue (SRW1) was recently described (YAMAGUCHI, publication cited above). The gene encoding CCS52Ms was cloned into the plasmid into pREP1 under the control of the *nmt1* promoter which is repressible by thiamine.

The cDNA of *ccs52Ms* obtained after cleavage of λ-ZAP (STRATAGENE) is digested with AgeI and partially with EcoRV. The AgeI-EcoRV fragment of 1.6 kb representing the coding region, with the exception of the first 4 codons, is cloned into a vector SKII

BLUESCRIPT (STRATAGENE) digested with XmaI (compatible with AgeI) and EcoRV. From this plasmid (pSK52B), the cDNA of *ccs52Ms* is cut by BamHI-EcoRV digestion and cloned into the BamHI-SmaI sites of the plasmid pREP1 [MAUNDRELL et al., Gene, 123, pp. 127-30, (1993)]. To generate an open reading frame in phase with the ATG codon for translation present in the vector under the control of the *nmtI* promoter, the DNA is digested with BamHI and the 5′ end is completed in the presence of the Klenow enzyme and of dNTPs. The religation of the blunt ends causes correct fusion, also verified by sequencing. This plasmid, called pREP52, is used to transform competent *S. pombe* SP-Q01 cells and the transformants are selected on EMM-thiamine agar plates, using the ESP kit (STRATAGENE). The vectors pREP1 [MAUNDRELL et al., Gene, 123, pp. 127-30, (1993)] and pESP1 (STRATAGENE) are used as negative controls; the positive control consists of *srw1* cloned into pREP1 [YAMAGUSHI et al., Mol. Biol. Cell., 8, pp. 2475-2486, (1997)].

The transformants of *S. pombe* SP-Q01 are cultured in 2 ml of 5 μM EMM-thiamine medium for 32 h at 30°C. The cells are washed twice with 10 ml of sterile water and resuspended in 5 ml of EMM medium. The cellular suspensions are divided into two halves: 2.5 ml are cultured with thiamine and 2.5 ml are cultured without thiamine, at 30°C. Culture aliquots are collected after 16 h and 24 h of culture and fixed with ethanol, stained with DAPI or with propidium iodide for analysis by flow cytometry and by microscopy [BEACH et al., Curr. Genet., 10, pp. 297-311, 1985)].

In the presence of thiamine, the expression of CCS52Ms is repressed and normal growth is observed.

In the absence of thiamine, the expression of CCS52Ms causes the inhibition of the growth of *S. pombe*, which is accompanied by endoreplication as illustrated in Figure 3B, which shows the presence of nuclei ≥ 4C, which is not observed in the control cells

of *S. pombe*, carrying the empty vector pREP1 (Figure 3A).

The morphology of the cells is also modified by the expression of CCS52Ms. A lengthening of the cells and an increase in the size of the nuclei are observed, which are identical to those observed during the expression of SRW1 [YAMAGUSHI et al., Mol. Biol. Cell., 8, pp. 2475-2486, (1997)], whereas no morphological change is observed when *S. pombe* carries only the vector pREP1.

In *S. pombe*, SRW1 is essential for the degradation of the mitotic cyclin CDC13. To verify if CCS52 acts in the same manner, the quantity of the CDC13 was evaluated in cultures of a strain (SY1) of *S. pombe*, carrying a deletion in the *srw1* gene, and not degrading CDC13.

The total proteins obtained from cultures of SY1 transformed with pREP1 (control) or with pREP1-*ccs52* was analysed by Western transfer, and visualized with the aid of anti-CDC13 antibodies.

In parallel, the expression of CDC2 kinase and that of α-tubulin were evaluated by visualization with the aid of anti-PSTAIR and anti-α-tubulin antibodies, respectively.

The results obtained show a very high reduction in CDC13 in the cells transformed with pREP1-*ccs52* compared with the control cells. By contrast, there is no variation in CDC2 and in α-tubulin.

These results confirm that CCS52 is a functional equivalent of SRW1.

EXAMPLE 4: PRODUCTION OF TRANSGENIC PLANTS TRANSFORMED WITH THE *CCS52MS* GENE

1. Expression of an antisense transcript and its action on the level of ploidy of *Medicago truncatula*.

In a first instance, the level of ploidy of various organs of *Medicago truncatula* (plant which is naturally diploid) was determined, by flow cytometry, in nontransformed plants.

The technique used is the following:

The nuclear DNA of freshly harvested plants is analysed by flow cytometry (EPICS V, Coulter), in accordance with the method of BROWN et al., (*A laboratory guide for Cellular and Molecular plant Biology*, 1991, 326-345, ed. Negrutiu et al., Birkhäuser, Basel), modified such that the nuclei are stained with DAPI at a final concentration of 5 µg/ml. The nuclear buffer I is used at 1% Triton X-100 for the nodules.

In young shoots, a quantity of DNA from 2C to 8C is found in the root and the cotyledon, whereas the hypocotyl also contains nuclei at 16C. In adult plants, the leaves are diploid, containing 95% of nuclei at 2C and 5% of nuclei at 4C. In the petioles and the nodules, nuclei from 2C to 32C were detected. However, the petiole contains predominantly nuclei at 2C, whereas the nodules contain predominantly nuclei at 4C.

An SstI-PvuII fragment of 1.2 kb containing 3/4 of the coding sequence of *ccs52Ms*, was placed in antisense orientation under the control of the 35S promoter, in a binary vector obtained from the vector pGPTV-BAR, carrying the *bar* gene for resistance to the herbicide BASTA as selectable marker, and multiple cloning sites. This construct is obtained by inserting the 35S promoter into a HindIII-XbaI fragment (obtained from pBI121, CLONTECH), into the HindIII-XbaI sites of the vector pGPTV-BAR. The *uidA* gene is then removed from the plasmid pGPTV-BAR by XbaI-SstI digestion at the level of the multiple cloning site.

To obtain the antisense construct of *ccs52Ms*, the SstI-PvuII fragment of 1.2 kb is cloned into the SmaI-SstI sites of the binary vector thus obtained.

These plasmids as well as a control plasmid, containing the *gus* gene instead of the antisense *ccs52* construct were introduced into *Agrobacterium tumefaciens* (EHA105) by electroporation and used to transform *Medicago truncatula* R108-1 according to the protocol described by HOFFMANN et al. [Mol. Plant

Microbe Interaction, 10, pp. 307-315, (1997)]; TRINH ET AL. [Plant Cell Reports, 17, pp. 345-355, (1998)].

The level of ploidy of the transgenic plants obtained was analysed, as described above and the level of endogenous transcripts was evaluated by RT-PCR. To differentiate the endogenous transcripts of *ccs52Mt* from the antisense transcripts, the pair of primers P55CL/P55CR is used for the endogenous transcripts and the pair of primers P55BL/P55CR for the antisense transcripts.

Sub B2

P55BL : TTTGGGGGTTGATGATTGTG
P55CL : CTCTCTACCGTTCTATCTCTTGGGA
P55CR : GGTAAAGATGCTACTTTGGTGGTGT

The position of these primers is schematically represented in Figure 4.

Figure 5A shows the results of evaluation of the quantity of endogenous *ccs52Mt* transcripts:

- by RT-PCR (□) in the transgenic lines A1, A3, A4, A7 and A32 and in the control plants containing the gus gene (C_{2n}), and
- by Northern transfer (■) in A4 and C_{2n} plants.

The results of analysis by flow cytometry are illustrated by Figure 5B, for the petioles of control plants containing the *gus* gene, diploids (C_{2n}) or tetraploids (C_{4n}), and of plants of the A4 line.

Out of 38 regenerated transgenic plants, 3 (A4, A7 and A32) showed a significantly reduced endoploidy, and in particular the plant A4. It is also in this line that the level of expression of the endogenous transcripts of *ccs52Ms* is the lowest, as shown in Figure 5B. The fact that a reduction in endoploidy was never observed before in other transgenic plants and are not observed in the control plants makes it possible to attribute this phenomenon to the impairment of the expression of CCS52Ms, and not to a secondary effect of transgenesis.

In addition, the plant A4 produces a quantity of seeds significantly less than that of the control plants. Moreover, it forms fewer side branches, and

develops only 2 nodules at the level of the roots, instead of the 50 nodules on average developed by the control plants cultured under the same conditions.

The impact of the partial suppression of the expression of *ccs52* on the development of the plant organs was also determined. For this purpose, the width of the petioles was measured and correlated with the percentages of endoreplicated nuclei (> 4C), in the T1 generation derived from the A4 line and in the C_{2n} and C_{4n} control plants.

The results are illustrated in Figure 6. Figure 6A which represents the width of the petiole as a function of the percentage of polyploid cells shows that, in the C_{2n} control plants (18 plants), the width of the petioles varies in correlation with the number of diploid cells. In the plants derived from A4 (36 plants), a more reduced variation in the size of the petioles and a lower percentage of polyploid cells are observed, which indicates that the degree of endoploidy can directly affect the final size of the plant organs.

12 of the 36 T1 plants derived from A4 contain less than 6% of endoreplicated nuclei (> 4C) in their petioles (Figure 6B). These plants [A4(s)] were grouped together and analysed separately from the rest of the A4 T1 plants [A4(w)] which exhibit less substantial phenotypic impairments.

Figure 6C shows that the width of the petioles in A4(w) plants is comparable to that of the diploid C_{2n} control plants; by contrast, the width of the petioles in the A4(s) plants is significantly less than that of the diploid C_{2n} control plants and the width of the petioles in the tetraploid C_{4n} control plants is significantly greater than that observed in the diploid plants.

The size of the leaves (which do not contain endoreplicated cells and whose endoploidy is not therefore affected by the level of expression of CCS52) was also measured. In this case, no significant difference is observed between the A4(w) plants, the

A4(s) plants and the diploid C_{2n} control plants; by contrast, the size of the leaves is significantly larger in the tetraploid C_{4n} control plants.

These results show that endoploidy affects the size of the plant organs, and that the modification of the expression of CCS52 acts at this level through a modification of the endoploidy.

<u>2. Expression of the CCS52Ms protein in transgenic plants.</u>

Expression vectors containing the *ccs5Ms* gene under the control of the 35S promoter, as well as expression vectors containing the *ccs52Ms* gene, under the control of a tissue-specific promoter, were constructed according to the following protocol:

For the tissue-specific expression of CCS52Ms, the cDNA is placed under the control of the *enod12AMs* and *Srglb3* promoters described by TRINH et al. [Plant Cell Reports, 17, pp. 345-355, (1998)], using as a vector pISV-BMCS, a derivative of pISV2301, and, instead of the complete *enod12AMs* promoter, only one 0.3 kb fragment thereof, considered to be sufficient for a nodule-specific expression [VIJN et al., Plant Mol. Biol., 28, pp. 1103-1110, (1995)].

Sub B3

Construction of pISV-BMCS: pISV2301 is digested with HindIII and SstI in order to eliminate the sequence of the 2X35S-AMV promoter, which is replaced by the following double-stranded BMCS oligonucleotide:

AGCTTCCCGGGGGAGCTCTAGACTCGAGCAGCT
AGGCCCCTCGAGATCTGAGCTCG.

This oligonucleotide contains the SmaI, SstI, XbaI and XhoI sites.

pISV-BMCS12A is constructed by cloning into pISV-BMCS of a fragment of the 0.3 kb of the *endo12AMs* promoter, obtained from the plasmid pPR89 [BAUER et al., Plant J., 10, pp. 91-105, (1996)].

pISV-BMCS-LB3 is constructed by digestion of pISV-BMCS with HindIII-SstI and cloning of a HindIII-SstI fragment containing the leghaemoglobin promoter of

Sesbania rostrata from pLP32 [TRINH et al., Plant Cell Reports, 17, pp. 345-355, (1998)].

These vectors were used to transform *Medicago truncatulata* according to the protocol described above for the antisense sequences.

During the regeneration of the transgenic plants, a significantly greater conversion of the calli to embryos is observed in plants transformed with the constructs expressing the *ccs52Ms* gene, than in plants transformed with the control construct, which indicates a positive effect of CCS52Ms on somatic embryogenesis.